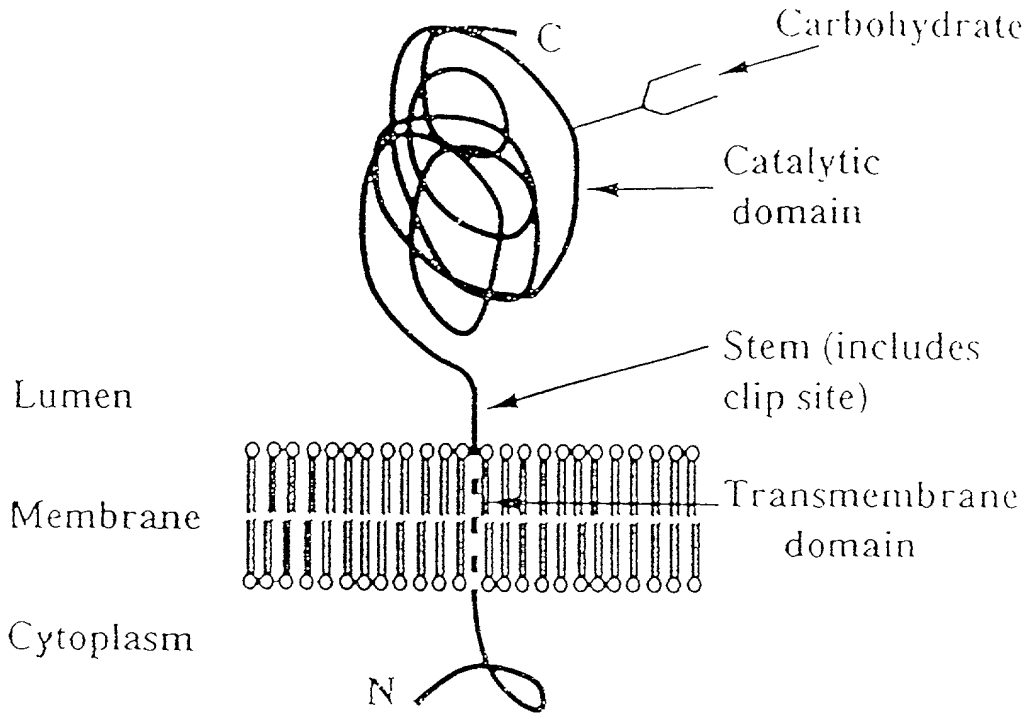

C
Carbohydrate
Catalytic
domain
Stem (includes
clip site)
Lumen
Membrane
Transmembrane
domain
Cytoplasm
N

Figure 1

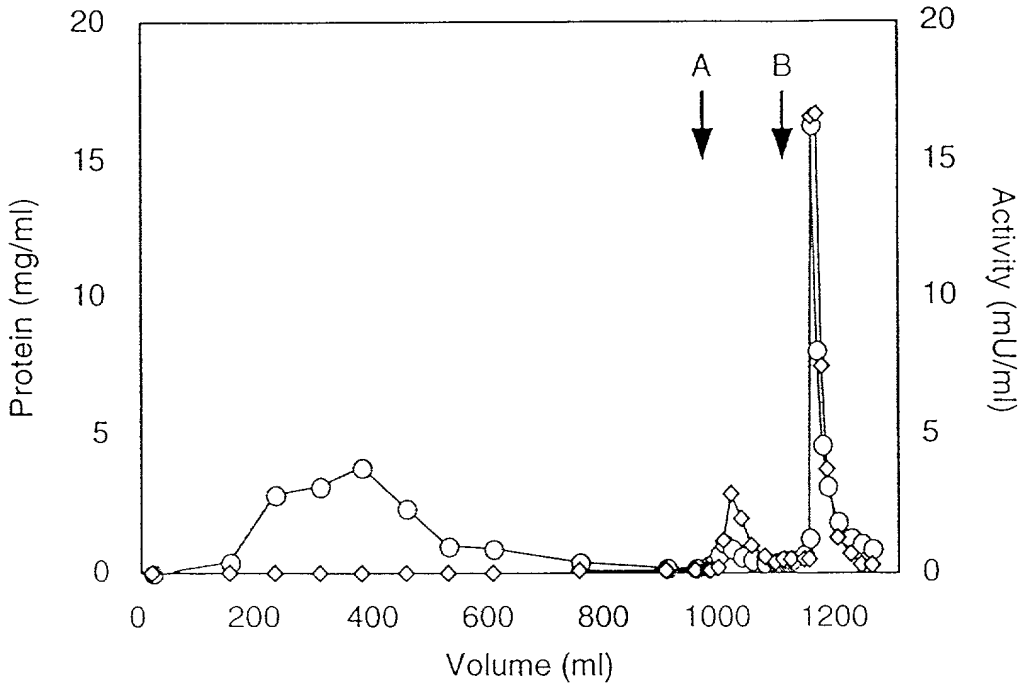

20
15
10
5
0
Protein (mg/ml)
A
B
20
15
10
5
0
Activity (mU/ml)
0
200
400
600
800
1000
1200
Volume (ml)

Figure 2

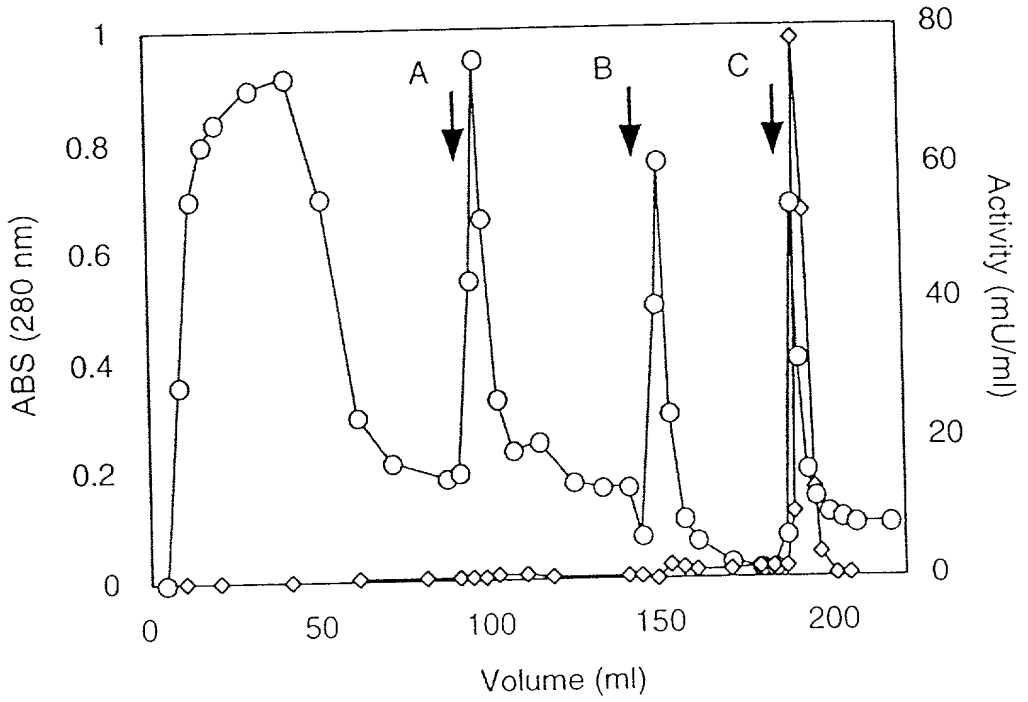

ABS (280 nm)
Activity (mU/ml)
Volume (ml)
A
B
C
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80
0
50
100
150
200

Figure 3

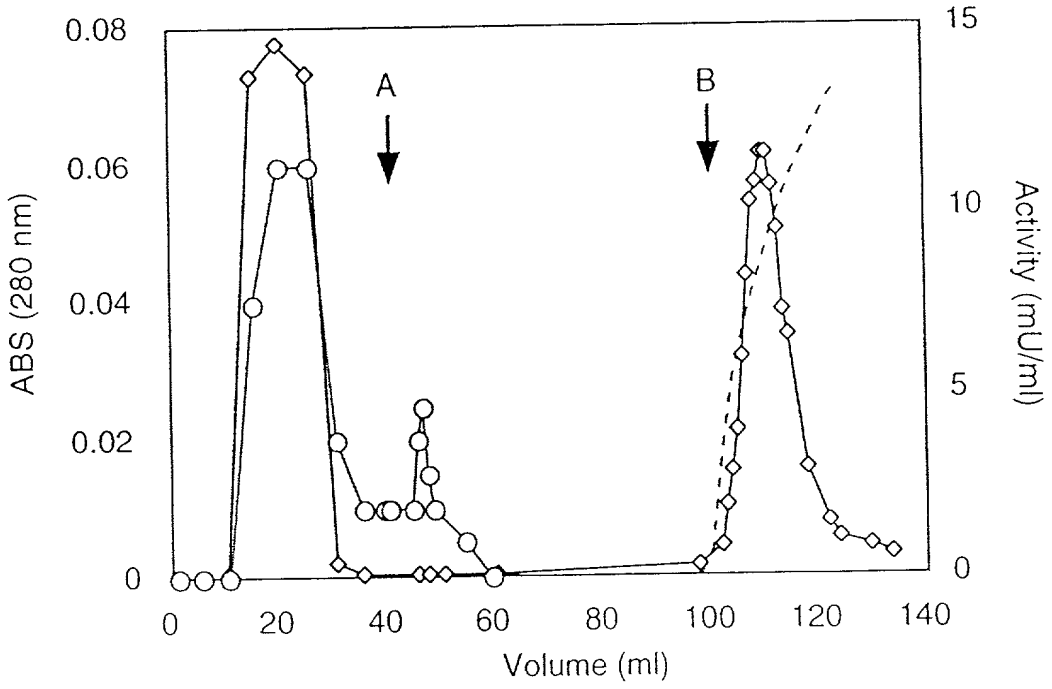

0.08
0.06
0.04
0.02
0
ABS (280 nm)
A
B
15
10
5
0
Activity (mU/ml)
0
20
40
60
80
100
120
140
Volume (ml)

Figure 4

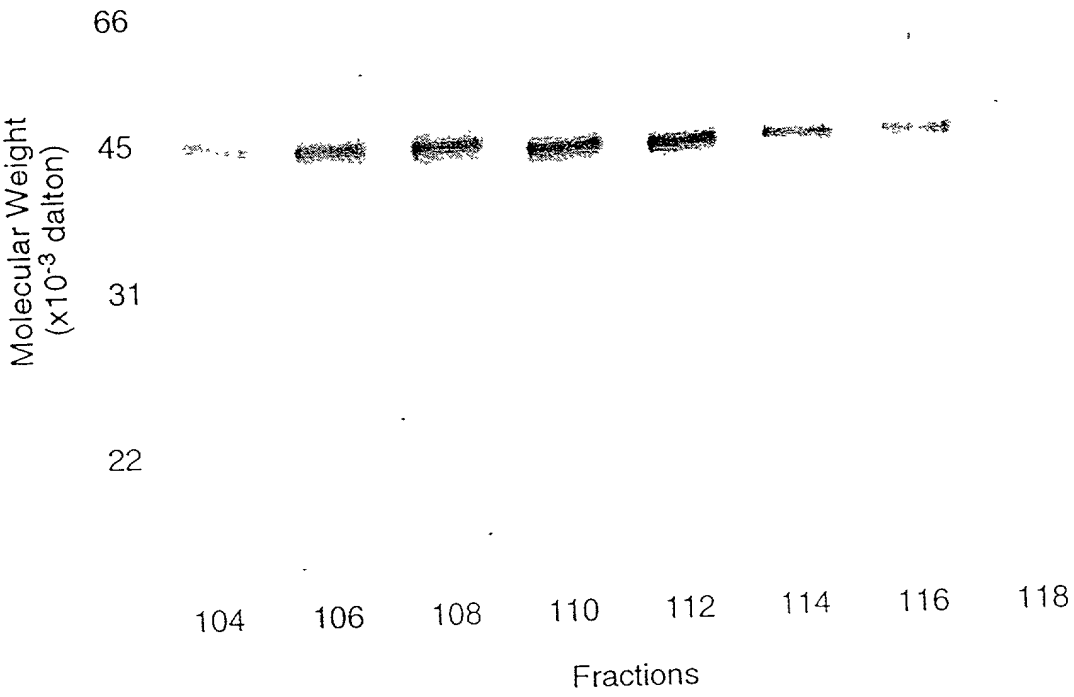

66
45
31
22
Molecular Weight
(x10-3 dalton)
104
106
108
110
112
114
116
118
Fractions

Figure 5

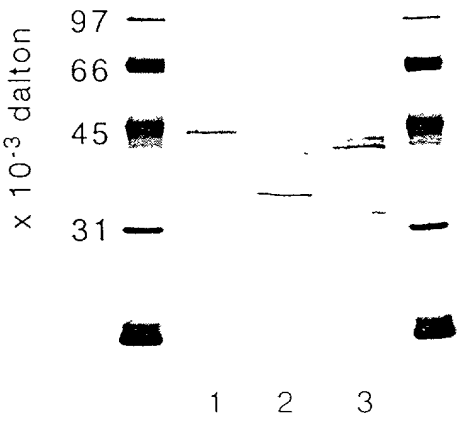


Figure 6

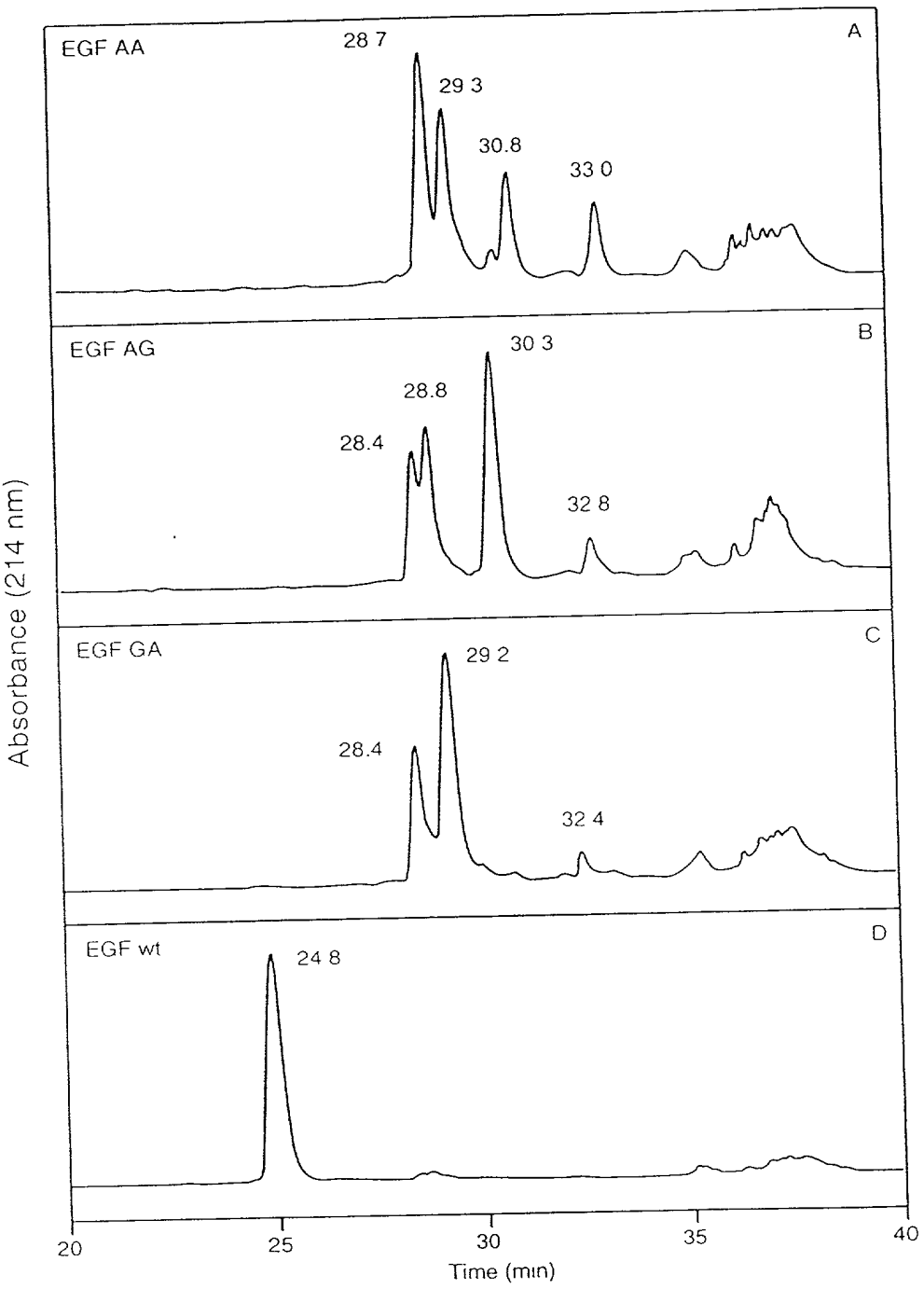

EGF AA
28 7
29 3
30.8
33 0
A
EGF AG
30 3
28.8
28.4
32 8
B
EGF GA
29 2
28.4
32 4
C
EGF wt
24 8
D
Absorbance (214 nm)
20
25
30
35
40
Time (min)

Figure 7

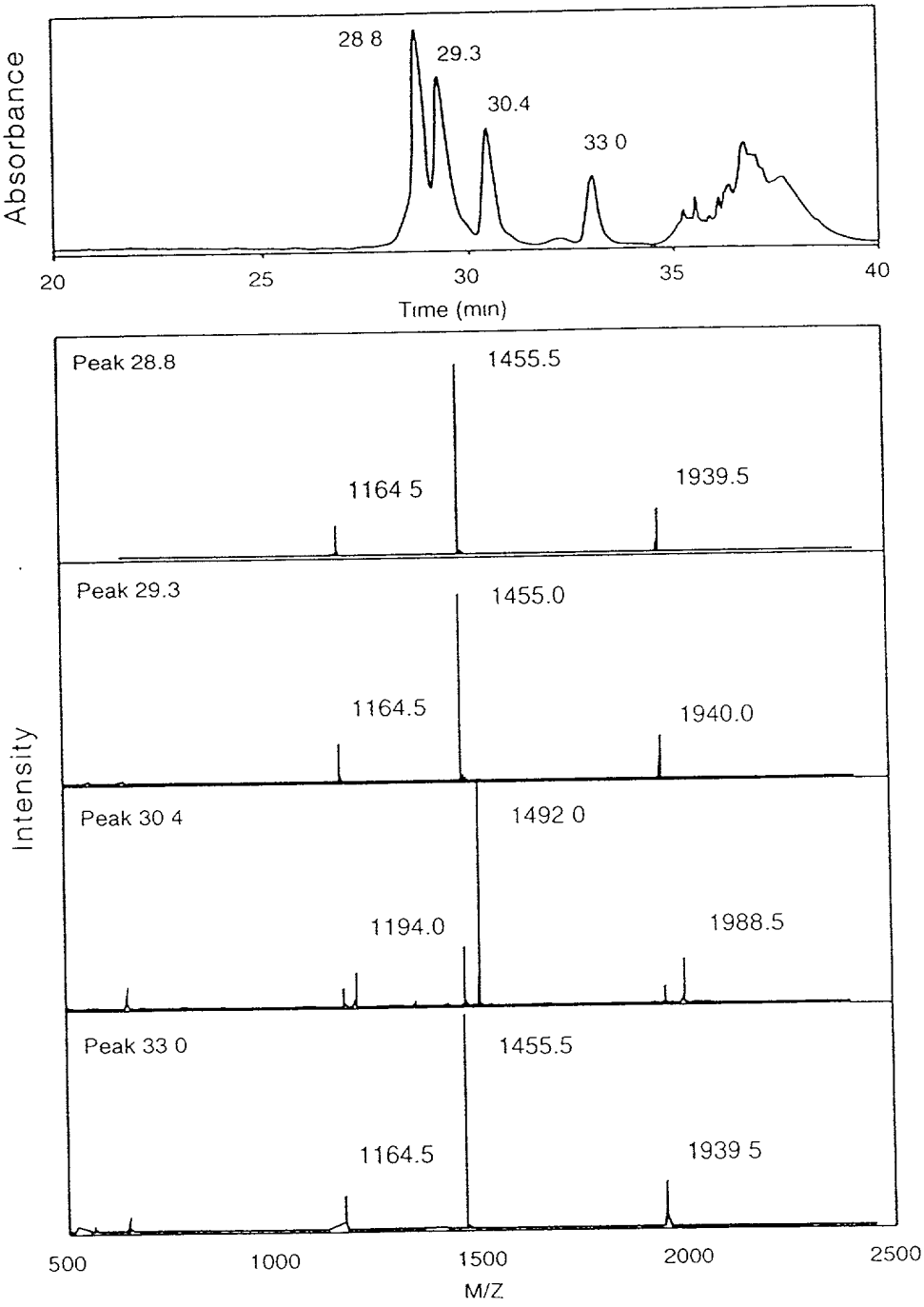

Absorbance
28 8
29.3
30.4
33 0
20
25
30
35
40
Time (min)
Peak 28.8
1455.5
1164 5
1939.5
Peak 29.3
1455.0
1164.5
1940.0
Peak 30 4
1492 0
1194.0
1988.5
Peak 33 0
1455.5
1164.5
1939 5
Intensity
500
1000
1500
2000
2500
M/Z

Figure 8

O-fucosyltransferase Sequence Similarity between Human, Hamster and C. elegans

```
C. Elegans    1  MSNYRYSKLNEEEISLEDMPSSANQILTRQEQIIQEQDDELELVGNSVRT

C. Elegans   51  LRGMSSMIGDELDQQSTMLDDLGQEMEYSETRLDTAMKKMAKLTHLEDGM

CHO           1  -----------------------------------------------RLA
C. Elegans  101  LLARRIVQSMQNDHGALSSPVFPRLCPSGLTTYVPYIVDFSSLTFHIFII

CHO           4  GSWDLAGTLLYXPXMGRFGNQADHFLGSLAFAKLXVRTLAVPPWIEYQHH
Human         1  -------------------NQADHFLGSLAFAKLLNRTLAVPPWIEYQHH
C. Elegans  151  IIIIIIDFCSQSQSKGRFGNQVDQFLGVLAFAKALDRTLVLPNFIEFKHP

CHO          54  KPPFTNLH------------------------------------------
Human        32  KPPFTNLHVSYQKYFKLEPLQAYHRVISLEDFMEKLAPTHWPPEKRVAYC
C. Elegans  201  ETKMIPFEFLFQVG----TVAKYTRVVTMQEFTKKIMPTHFVGTPRQAIY

Human        82  FEVAAQRSPDKKTCPMKEGNPFGPFWDQFHVSFNKSELFTGISFSASYRE
C. Elegans  247  DKSAEPGCHSK------EGNPFGPYWDQIDVSFVGDEYFGDIPGGFDLNQ

Human       132  QWSQR-----FSPKEHPVLALPGAPAQFPVLEEHRPLQKYMVWSDEMVKT
C. Elegans  291  MGSRKKWLEKFPSEEYPVLAFSSAPAPFPSKGKVWSIQKYLRWSSRITEQ

Human       177  GEAQIHAHLVRPYVGIHLRIGSDWKNACAMLKDGTAGSHFMASPQCVGYS
C. Elegans  341  AKKFISANLAKPFVAVHLRNDADWVRVCEHIDTTTNRPLFASEQCLG---

Human       227  RSTAAPLTMTMCLPDLKEIQRAVKLWVRSLDAQSVYVATDSESYVPELQQ
C. Elegans  388  --------------------------------EGHHLGTLTKEICSPSKQ

Human       277  LFKGKVKVVSLKPEVAQVDLYILGQADHFIGNCVSSFTAFVKRERDLQGR
C. Elegans  406  QILEQIEAHRQEPDDMYTSLAIMGRADLFVGNCVSTFSHIVKRERDHAGQ

Human       327  PSSFFGMDRPPKLRDEF--
C. Elegans  456  SPRPSAFFGIRAVKRHIDL
```

Similarity between hamster O-fucosyltransferase and human and C elegans genes. N-terminal peptide sequence of hamster O-fucosyltransferase is shaded. Human sequence is a partial cDNA of unknown protein from a myeloblast cell line and C. elegans gene is a computer generated coding sequence from its genome.

Figure 9

Northern Blot For O-Fucosyltransferase

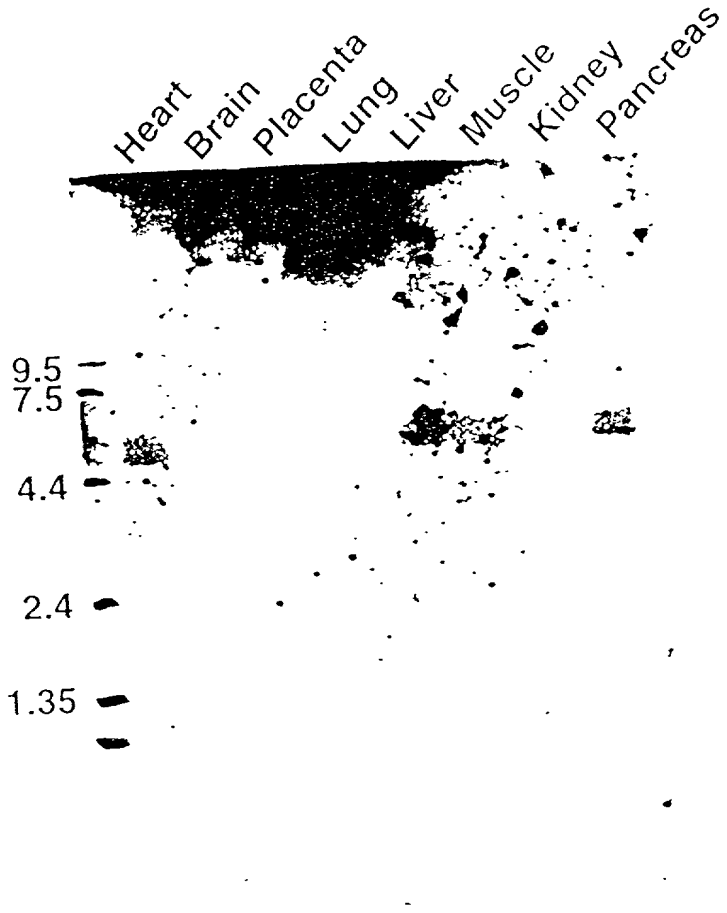
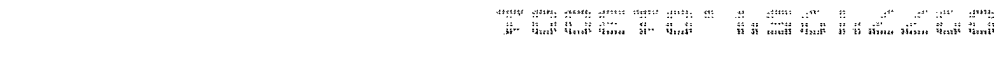


Figure 10



GAACCAGGCC GATCACTTCT TGGGCTCTCT GGCATTTGCA AGCTGCTAA ACCGTACCTT GGCTGTCCCT CCTTGGATTG AGTACCAGCA TCACAAGCCT CCTTTCACCA ACCTCCATGT
CTTGGTCCGG CTAGTGAAGA ACCCGAGAGA CCGTAAACGT TTCGACGATT TGGCATGGAA CCGACAGGGA GGAACCTAAC TCATGGTCGT AGTGTTCGGA GGAAAGTGGT TGGAGGTACA

GTCCTACCAG AAGTACTTCA AGCTGGAGCC CCTCCAGGCT TACCATCGGG TCATCAGCTT GGAGGATTTC ATGGAGAAGC TGGCACCCAC CCACTGGCCC CCTGAGAAGC GGGTGGCATA
CAGGATGGTC TTCATGAAGT TCGACCTCGG GGAGGTCCGA ATGGTAGCCC AGTAGTCGAA CCTCCTAAAG TACCTCTTCG ACCGTGGGTG GGTGACCGGG GGACTCTTCG CCCACCGTAT

CTGCTTTGAG GTGGCAGCCC AGCGAAGCCC AGATAAGAAG ACGTGCCCCA TGAAGGAAGG AAACCCCTTT GGCCCATTCT GGGATCAGTT TCATGTGAGT TTCAACAAGT CGGAGCTTTT
GACGAAACTC CACCGTCGGG TCGCTTCGGG TCTATTCTTC TGCACGGGGT ACTTCCTTCC TTTGGGGAAA CCGGGTAAGA CCCTAGTCAA AGTACACTCA AAGTTGTTCA GCCTCGAAAA

TACAGGCATT TCCTTCAGTG CTTCCTACAG AGAACAATGG AGCCAGAGAT TTTCTCCAAA GGAACATCCG GTGCTTGCCC TGCCAGGAGC CCCAGCCCAG TTCCCCGTCC TAGAGGAACA
ATGTCCGTAA AGGAAGTCAC GAAGGATGTC TCTTGTTACC TCGGTCTCTA AAAGAGGTTT CCTTGTAGGC CACGAACGGG ACGGTCCTCG GGGTCGGGTC AAGGGGCAGG ATCTCCTTGT

CAGGCCACTA CAGAAGTACA TGGTATGGTC AGACGAAATG GTGAAGACGG GAGAGGCCCA GATTCATGCC CACCTTGTCC GGCCCTATGT GGGCATTCAT CTGCGCATTG GCTCTGACTG
GTCCGGTGAT GTCTTCATGT ACCATACCAG TCTGCTTTAC CACTTCTGCC CTCTCCGGGT CTAAGTACGG GTGGAACAGG CCGGGATACA CCCGTAAGTA GACGCGTAAC CGAGACTGAC

GAAGAACGCC TGTGCCATGC TGAAGGACGG GACTGCAGGC TCGCACTTCA TGGCCTCTCC GCAGTGTGTG GGCTACAGCC GCAGCACAGC GGCCCCCCTC ACGATGACTA TGTGCCTGCC
CTTCTTGCGG ACACGGTACG ACTTCCTGCC CTGACGTCCG AGCGTGAAGT ACCGGAGAGG CGTCACACAC CCGATGTCGG CGTCGTGTCG CCGGGGGGAG TGCTACTGAT ACACGGACGG

TGACCTGAAG GAGATCCAGA GGGCTGTGAA GCTCTGGGTG AGGTCGCTGG ATGCCCAGTC GGTCTACGTT GCTACTGATT CCGAGAGTTA TGTGCCTGAG CTCCAACAGC TCTTCAAAGG
ACTGGACTTC CTCTAGGTCT CCCGACACTT CGAGACCCAC TCCAGCGACC TACGGGTCAG CCAGATGCAA CGATGACTAA GGCTCTCAAT ACACGGACTC GAGGTTGTCG AGAAGTTTCC

GAAGGTGAAG GTGGTGAGCC TGAAGCCTGA GGTGGCCCAG GTCGACCTGT ACATCCTCGG CCAAGCCGAC CACTTTATTG GCAACTGTGT CTCCTCCTTC ACTGCCTTTG TGAAGCGGGA
CTTCCACTTC CACCACTCGG ACTTCGGACT CCACCGGGTC CAGCTGGACA TGTAGGAGCC GGTTCGGCTG GTGAAATAAC CGTTGACACA GAGGAGGAAG TGACGGAAAC ACTTCGCCCT

GCGGGACCTC CAGGGGAGGC CGTCTTCTTT CTTCGGCATG GACAGGCCCC CTAAGCTGCG GGACGAGTTC TGATTCTGGC CGGAGCACCA GACCCTCTGA TCCTGGAGGG ACCAGAGTCT
CGCCCTGGAG GTCCCCTCCG GCAGAAGAAA GAAGCCGTAC CTGTCCGGGG GATTCGACGC CCTGCTCAAG ACTAAGACCG GCCTCGTGGT CTGGGAGACT AGGACCTCCC TGGTCTCAGA

GAGCTGGTCC TTCCAGCCAG GCCTGGCAGC CAGAGGTGCT CCGGGATTGC AAACTCCTCT TCTCACCTGC CAAAGATGGA GAAGAGTGCC AGGGACCCCT CAAGGAGGGA GACGCTCCAT
CTCGACCAGG AAGGTCGGTC CGGACCGTCG GTCTCCACGA GGCCCTAACG TTTGAGGAGA AGAGTGGACG GTTTCTACCT CTTCTCACGG TCCCTGGGGA GTTCCTCCCT CTGCGAGGTA

ATCCCAGGGC ATAGGACTTG CAGGTTCCTA GGAGCAGGAG CATCTCCCAT CGCACGTGCT TTCTGCTCTT CTGGGAATTT CTCACACTGG CAAAGCAGTC CAGCCTCCGT CTTCTGGTCC
TAGGGTCCCG TATCCTGAAC GTCCAAGGAT CCTCGTCCTC GTAGAGGGTA GCGTGCACGA AAGACGAGAA GACCCTTAAA GAGTGTGACC GTTTCGTCAG GTCGGAGGCA GAAGACCAGG

ACTCTGCTCT GAGCAGCCTG GGATGCTGAA CTCTTCAGAG AGATTTTTTT ATAGAGAGAT TTCTATAATT TTGATACAAG GTCATGACTA TCCTAGAACT CTCTGTGGTT TTTGAAAATC
TGAGACGAGA CTCGTCGGAC CCTACGACTT GAGAAGTCTC TCTAAAAAAA TATCTCTCTA AAGATATTAA AACTATGTTC CAGTACTGAT AGGATCTTGA GAGACACCAA AAACTTTTAG

ATTGAATTC
TAACTTAAG

Human KIAA0180 First EcoRI Fragment. The first Eco RI fragment of the cDNA contains a partial coding sequence within complete amino terminus. The region which matched with CHO peptide sequence is shaded. The two oligonucleotides used to make the probe for the northern blot (Figure 2) are over-scored and double-underlined. The nucleotides over-scored and underlined are two primers used in PCR reaction as described in Methods.

Figure 11

Figure 12 A

```
   1  ATGCCCGCGG GCTCCTGGGA CCCGGCCGGT TACCTGCTCT ACTGCCCCTG CATGGGGCGC TTTGGGAACC AGGCCGATCA CTTCTTGGGC TCTCTGGCAT
   1  M  P  A  G  S  W  D  P  A  G  Y  L  L  Y  C  P  C  M  G  R  F  G  N  Q  A  D  H  F  L  G  S  L  A  F

 101  TTGCAAAGCT GCTAAACCGT ACCTTGGCTG TCCCTCCTTG GATTGAGTAC CAGCATCACA AGCCTCCTTT CACCAACCTC CATGTGTCCT ACCAGAAGTA
  35    A  K  L  L  N  R  T  L  A  V  P  P  W  I  E  Y  Q  H  H  K  P  P  F  T  N  L  H  V  S  Y  Q  K  Y

 201  CTTCAAGCTG GAGCCCCTCC AGGCTTACCA TCGGGTCATC AGCTTGGAGG ATTTCATGGA GAAGCTGGCA CCCACCCACT GGCCCCCTGA GAAGCGGGTG
  68    F  K  L  E  P  L  Q  A  Y  H  R  V  I  S  L  E  D  F  M  E  K  L  A  P  T  H  W  P  P  E  K  R  V

 301  GCATACTGCT TTGAGGTGGC AGCCCAGCGA AGCCCAGATA AGAAGACGTG CCCCATGAAG GAAGGAAACC CCTTTGGCCC ATTCTGGGAT CAGTTTCATG
 101  A  Y  C  F  E  V  A  A  Q  R  S  P  D  K  K  T  C  P  M  K  E  G  N  P  F  G  P  F  W  D  Q  F  H  V

 401  TGAGTTTCAA CAAGTCGGAG CTTTTTACAG GCATTTCCTT CAGTGCTTCC TACAGAGAAC AATGGAGCCA GAGATTTTCT CCAAAGGAAC ATCCGGTGCT
 135    S  F  N  K  S  E  L  F  T  G  I  S  F  S  A  S  Y  R  E  Q  W  S  Q  R  F  S  P  K  E  H  P  V  L

 501  TGCCCTGCCA GGAGCCCCAG CCCAGTTCCC CGTCCTAGAA GAACACAGGC CACTACAGAA GTACATGGTA TGGTCAGACG AAATGGTGAA GACGGGAGAG
 168    A  L  P  G  A  P  A  Q  F  P  V  L  E  E  H  R  P  L  Q  K  Y  M  V  W  S  D  E  M  V  K  T  G  E

 601  GCCCAGATTC ATGCCCACCT TGTCCGGCCC TATGTGGGCA TTCATCTGCG CATTGGCTCT GACTGGAAGA ACGCCTGTGC CATGCTGAAG GACGGGACTG
 201  A  Q  I  H  A  H  L  V  R  P  Y  V  G  I  H  L  R  I  G  S  D  W  K  N  A  C  A  M  L  K  D  G  T  A

 701  CAGGCTCGCA CTTCATGGCC TCTCCGCAGT GTGTGGGCTA CAGCCGCAGC ACAGCGGCCC CCCTCACGAT GACTATGTGC CTGCCTGACC TGAAGGAGAT
 235    G  S  H  F  M  A  S  P  Q  C  V  G  Y  S  R  S  T  A  A  P  L  T  M  T  M  C  L  P  D  L  K  E  I

 801  CCAGAGGGCT GTGAAGCTCT GGGTGAGGTC GCTGGATGCC CAGTCGGTCT ACGTTGCTAC TGATTCCGAG AGTTATGTGC CTGAGCTCCA ACAGCTCTTC
 268  Q  R  A  V  K  L  W  V  R  S  L  D  A  Q  S  V  Y  V  A  T  D  S  E  S  Y  V  P  E  L  Q  Q  L  F

 901  AAAGGGAAGG TGAAGGTGGT GAGCCTGAAG CCTGAGGTGG CCCAGGTCGA CCTGTACATC CTCGGCCAAG CCGACCACTT TATTGGCAAC TGTGTCTCCT
 301  K  G  K  V  K  V  V  S  L  K  P  E  V  A  Q  V  D  L  Y  I  L  G  Q  A  D  H  F  I  G  N  C  V  S  S

1001  CCTTCACTGC CTTTGTGAAG CGGGAGCGGG ACCTCCAGGG GAGGCCGTCT TCTTTCTTCG GCATGGACAG GCCCCCTAAG CTGCGGGACG AGTTCTGATT
 335    F  T  A  F  V  K  R  E  R  D  L  Q  G  R  P  S  S  F  F  G  M  D  R  P  P  K  L  R  D  E  F  O

1101  CTGGCCGGAG CACCAGACCC TCTGATCCTG GAGGGACCAG AGTCTGAGCT GGTCCTTCCA GCCAGGCCTG GCAGCCAGAG GTGCTCCGGG ATTGCAAACT

1201  CCTCTTCTCA CCTGCCAAAG ATGGAGAAGA GTGCCAGGGA CCCCTCAAGG AGGGAGACGC TCCATATCCC AGGGCATAGG ACTTGCAGGT TCCTAGGAGC

1301  AGGAGCATCT CCCATCGCAC GTGCTTTCTG CTCTTCTGGG AATTTCTCAC ACTGGCAAAG CAGTCCAGCC TCCGTCTTCT GGTCCACTCT GCTCTGAGCA

1401  GCCTGGGATG CTGAACTCTT CAGAGAGATT TTTTTATAGA GAGATTTCTA TAATTTTGAT ACAAGGTCAT GACTATCCTA GAACTCTCTG TGGTTTTTGA

1501  AAATCATTGA ATTC
```

Figure 12 B

```
                              10        20        30        40        50        60
Human      MPAGSWDPAGYLLYCPCMGRFGNQADHFLGSLAFAKLLNRTLAVPPWIEYQHHKPPFTNLH
                 ***** ****** * ********************  **********************
CHO        RLAGSWDLAGYLLYXPXMGRFGNQADHFLGSLAFAKLXVRTLAVPPWIEYQHHKPPFTNLH
                              10        20        30        40        50        60
```

Human hear O-fucosyltransferase Sequence. Upper panel, compiled sequence from positive cDNA clones. The region that matches with CHO cell sequence is shaded The residue A at position 540 of the DNA sequence is different from that of human KIAA0180 (G at position 475 of Figure X). The peptide sequences are the same. Lower panel, comparison of O-fucosyltransferase amino terminal sequences from human heart and CHO cells.

Figure 13 A

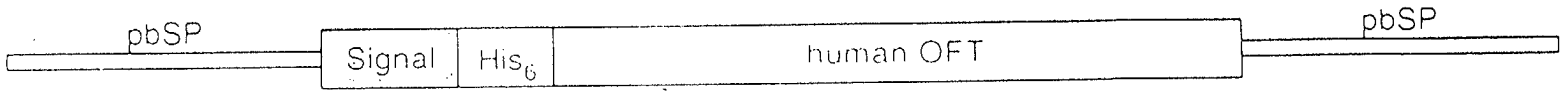


Figure 13 B

```
4101  TTATTCATAC CGTCCCACCA TCGGGCGCGG ATCAGATCCA TGGCCAAGTT CCTGGTCAAC GTGGCCCTGC TGCTGCTGCT GCTGCTGCTG TCCGGAGCCT
   1                                                M  A  K  F  L  V  N  V  A  L  L  L  L  L  L  L  S  G  A  W

4201  GGGCCCATAT GAGATCCCAT CACCATCACC ATCACATGCC CGCGGGCTCC TGGGACCCGG CCGGTTACCT GCTCTACTGC CCCTGCATGG GGCGCTTTGG
  22    A  H  M  R  S  H  H  H  H  H  H  M  P  A  G  S  W  D  P  A  G  Y  L  L  Y  C  P  C  M  G  R  F  G

4301  GAACCAGGCC GATCACTTCT TGGGCTCTCT GGCATTTGCA AAGCTGCTAA ACCGTACCTT GGCTGTCCCT CCTTGGATTG AGTACCAGCA TCACAAGCCT
  55    N  Q  A  D  H  F  L  G  S  L  A  F  A  K  L  L  N  R  T  L  A  V  P  P  W  I  E  Y  Q  H  H  K  P

4401  CCTTTCACCA ACCTCCATGT GTCCTACCAG AAGTACTTCA AGCTGGAGCC CCTCCAGGCT TACCATCGGG TCATCAGCTT GGAGGATTTC ATGGAGAAGC
  88    P  F  T  N  L  H  V  S  Y  Q  K  Y  F  K  L  E  P  L  Q  A  Y  H  R  V  I  S  L  E  D  F  M  E  K  L

4501  TGGCACCCAC CCACTGGCCC CCTGAGAAGC GGGTGGCATA CTGCTTTGAG GTGGCAGCCC AGCGAAGCCC AGATAAGAAG ACGTGCCCCA TGAAGGAAGG
 122    A  P  T  H  W  P  P  E  K  R  V  A  Y  C  F  E  V  A  A  Q  R  S  P  D  K  K  T  C  P  M  K  E  G

4601  AAACCCCTTT GGCCCATTCT GGGATCAGTT TCATGTGAGT TTCAACAAGT CGGAGCTTTT TACAGGCATT TCCTTCAGTG CTTCCTACAG AGAACAATGG
 155    N  P  F  G  P  F  W  D  Q  F  H  V  S  F  N  K  S  E  L  F  T  G  I  S  F  S  A  S  Y  R  E  Q  W

4701  AGCCAGAGAT TTTCTCCAAA GGAACATCCG GTGCTTGCCC TGCCAGGAGC CCCAGCCCAG TTCCCCGTCC TAGAGGAACA CAGGCCACTA CAGAAGTACA
 188    S  Q  R  F  S  P  K  E  H  P  V  L  A  L  P  G  A  P  A  Q  F  P  V  L  E  E  H  R  P  L  Q  K  Y  M

4801  TGGTATGGTC AGACGAAATG GTGAAGACGG GAGAGGCCCA GATTCATGCC CACCTTGTCC GGCCCTATGT GGGCATTCAT CTGCGCATTG GCTCTGACTG
 222    V  W  S  D  E  M  V  K  T  G  E  A  Q  I  H  A  H  L  V  R  P  Y  V  G  I  H  L  R  I  G  S  D  W

4901  GAAGAACGCC TGTGCCATGC TGAAGGACGG GACTGCAGGC TCGCACTTCA TGGCCTCTCC GCAGTGTGTG GGCTACAGCC GCAGCACAGC GGCCCCCCTC
 255    K  N  A  C  A  M  L  K  D  G  T  A  G  S  H  F  M  A  S  P  Q  C  V  G  Y  S  R  S  T  A  A  P  L

5001  ACGATGACTA TGTGCCTGCC TGACCTGAAG GAGATCCAGA GGGCTGTGAA GCTCTGGGTG AGGTCGCTGG ATGCCCAGTC GGTCTACGTT GCTACTGATT
 288    T  M  T  M  C  L  P  D  L  K  E  I  Q  R  A  V  K  L  W  V  R  S  L  D  A  Q  S  V  Y  V  A  T  D  S

5101  CCGAGAGTTA TGTGCCTGAG CTCCAACAGC TCTTCAAAGG GAAGGTGAAG GTGGTGAGCC TGAAGCCTGA GGTGGCCCAG GTCGACCTGT ACATCCTCGG
 322    E  S  Y  V  P  E  L  Q  Q  L  F  K  G  K  V  K  V  V  S  L  K  P  E  V  A  Q  V  D  L  Y  I  L  G

5201  CCAAGCCGAC CACTTTATTG GCAACTGTGT CTCCTCCTTC ACTGCCTTTG TGAAGCGGGA GCGGGACCTC CAGGGGAGGC CGTCTTCTTT CTTCGGCATG
 355    Q  A  D  H  F  I  G  N  C  V  S  S  F  T  A  F  V  K  R  E  R  D  L  Q  G  R  P  S  S  F  F  G  M

5301  GACAGGCCCC CTAAGCTGCG GGACGAGTTC TGATTCTGGC CGGAGCACCA GACCCTCTGA TCCTGGAGGG ACCAGAGTCT GAGCTGGTCC TTCCAGCCAG
 388    D  R  P  P  K  L  R  D  E  F  0
```

Plasmid construct for expression of human)-fucosyltransferase Upper panel is a schematic drawing of the plasmid Lower panel is the sequence of the insert. The artificial signal peptide is shaded and the poly histidine tag is double underlined Human hear O-fucosyltransferase part is the same as in Figure 5.

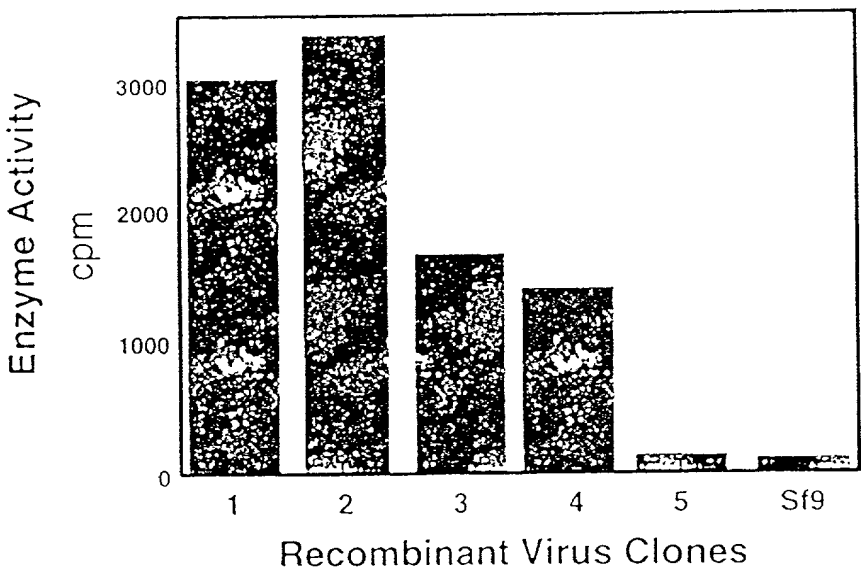

Enzyme Activity
cpm
3000
2000
1000
0
1
2
3
4
5
Sf9
Recombinant Virus Clones

Figure 14

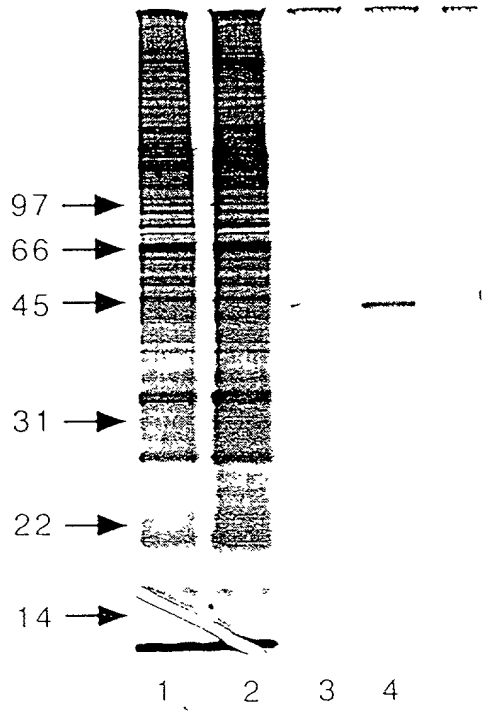


Figure 15